FACE MOTION INFLUENCES EYE MOVEMENT PATTERNS
AND FACE PROCESSING IN CHILDREN

by

Steve Perrotta

A thesis submitted in conformity with the requirements
for the degree of Master of Arts
Graduate Department of Applied Psychology and Human Development
Ontario Institute for Studies in Education
University of Toronto

© Copyright by Steve Perrotta (2014)
Face Motion Influences Eye Movement Patterns and Face Processing in Children  
Master of Arts in Developmental Psychology & Education, 2014  
Steve Perrotta  
Department of Applied Psychology and Human Development  
University of Toronto

Abstract

The present study examined the influence of face motion on processing strategy and eye movement patterns in 4- to 6-year-old children. The results indicated that face motion influenced featural processing similar to previous studies with adults, though this effect was not different from that of static images. Consistent with previous studies in infancy and childhood, eye-tracking analyses revealed that observing moving faces attracted longer fixation times in the mouth region and less at the eye regions compared to static images. These findings suggest that, while face processing strategy was no different when shown in motion compared to static, face motion influences eye movement patterns in early childhood.
ACKNOWLEDGEMENTS

I would like to thank my lab supervisor, Dr. Kang Lee, for his generosity and support. His feedback enabled me to approach this investigation in a straightforward manner. I am indebted to my thesis supervisor, Naiqi (Gabriel) Xiao, for his strong insight about the data, for teaching me to work efficiently with Matlab, R, and Excel, and for his encouragement throughout this process. I am also thankful for the members of the Early Social Development Lab and the Child Development Research Group, especially to Sarah Zanette and Reem Tawfik, for their unwavering support when the going got tough.

I would like to thank several professors from my alma mater, Nipissing University, whom have stayed in touch and offered their wisdom about graduate studies: Drs. Deborah Flynn, Jeff Nicol, Peter Chow, and Dana Murphy. Also, a big thank-you goes to my mother, Fran Perrotta, and my father, Mario Perrotta, for their love and support.

Finally, I would like to thank all participants involved in this study who gave their time to make this research possible.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>iii</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>iv</td>
</tr>
<tr>
<td>List of Figures</td>
<td>v</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>METHODS</td>
<td>15</td>
</tr>
<tr>
<td>RESULTS</td>
<td>19</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>21</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>27</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1. Illustrations of the elastic motion (dynamic) and face image (static) target face conditions.

Figure 2. Examples of composite faces. The upper panel face pair is misaligned, and the lower panel face pair is aligned. One face from the pair is the target face, and the other is the foil face.

Figure 3. An example of areas of interest (AOIs) depicted in the familiarization phase: left eye, right eye, nose, mouth, upper half, lower half.

Figure 4. Mean accuracy of target composite face recognition.

Figure 5. Proportion of fixation time on facial features by target face type.
Introduction

Faces may be the most important type of visual stimuli in our environment, as we are likely to encounter them every day. The human face provides observers with a wealth of characteristic information, such as gender, age, and race. Furthermore, eye gaze and emotional expressions provide important information in social contexts. One well-documented area of face perception research is concerned with how this information influences our ability to process faces. Through our experience with observing faces, humans may develop advanced capabilities for processing faces. Face recognition, a skill in which one matches the identity information of an individual face to that encoded in memory, is a quality that can be measured to examine how faces are processed.

Face recognition abilities begin developing at birth. Newborns can distinguish between a face-like schematic and a face with scrambled features (Goren, Sarty, & Wu, 1975), and after just a few days they can recognize their mother’s face (Bushnell, Sai, & Mullin, 1989). Despite this, there still exists some controversy on the nature of humans’ face processing abilities; while some evidence suggests faces are a special class of stimuli of which we are intuitive, other research suggests that our superior face processing abilities are reflective of visual learning due to considerable exposure to faces (Diamond & Carey, 1986; Kanwisher, 2000). Nevertheless, researchers generally agree that adults’ ability for advanced face processing is due to their experience with perceiving faces (Pascalis et al., 2011).

Support from prior investigations strongly suggests that, throughout development, there exists a qualitative shift in the primary type of processing used for recognition (Carey & Diamond, 1977, 1994; Schwarzer, Zauner, & Jovanovic, 2007). Some
researchers attribute this idea to the role of experience in shaping children’s face processing expertise (see Lee et al., 2011, for a review). Evidently, a face is not recognized simply from its individual features, but by combining the configuration of these features into a whole face (Galton, 1879). For example, a friend’s face is not only identified by the shape of their eyes, nose, and mouth, but rather the spatial combination of these features that create an instant and distinctive representation of that person. It would follow, then, that various types of facial information combine to perceive a face.

Researchers generally agree that faces are processed in ways that are qualitatively distinct from each other (Tanaka & Farah, 1993; Young, Hellawell, & Hay, 1987). One manner that faces are processed is featurally, in which one tends to process facial features (e.g., eyes, nose, and mouth) individually. Another manner in which faces are processed is holistically, in which one tends to integrate facial features as gestalt, or an unbroken whole (see Tanaka & Gordon, 2011, for a review). Faces are poorly recognized when shown inverted rather than upright (Yin, 1969) – this is believed to be due to a disruption of holistic face information, while featural information stays relatively intact (Freire, Lee, & Symons, 2000; Richler, Cheung, & Gauthier, 2011). However, researchers have found that inversion disrupts not only holistic processing, but processing of featural information as well (Michel, Rossion, Han, Chung, & Caldara, 2006; Yovel & Kanwisher, 2004). This has raised concern about whether the face-inversion paradigm is an appropriate test of holistic processing. More recently, the composite face effect (CFE) paradigm has been considered a more reliable and direct measure of holistic face processing, and is more widely used in such investigations (e.g., McKone, 2008; Mondloch, Pathman, Maurer, Le Grand, & de Schonen, 2007; Young et al., 1987). A composite face is comprised of upper
and lower face halves that belong to two different individuals. The composite face halves may be aligned to form a whole face, or misaligned so that the face halves are perceived as separate. The composite face effect is revealed when observers have more difficulty recognizing the identity of one of the face halves when they are aligned than when they are misaligned. The idea behind this effect is that holistic processing will interfere with identification of one face half from the other if they are aligned to form a face gestalt; likewise, this interference is reduced when the two face halves are misaligned, since the halves will not be perceived as a whole face. Therefore, there is consensus that the observation of a composite face effect indicates the presence of holistic processing, while a lack of this effect suggests the presence of featural processing.

Previous literature generally shows that those with an immature face processing system (i.e., infants and young children) predominantly use featural processing – they rely on the appearance of facial features for recognition (McKone, 2008; Mondloch, Geldart, Maurer, & Le Grand, 2003). That being said, evidence of holistic processing abilities is present in infancy (Cohen & Cashon, 2001). For example, before one year of age, infants learn to process the eyes and mouth in relation to the face (Schwarzer et al., 2007; Schwarzer and Zauner, 2003). Moreover, Ferguson, Kulkofsky, Cashon, & Casasola (2009) demonstrated that infants as young as 4 months can process facial features as a unified whole. While faces may be processed holistically to some degree in infancy and early childhood, it is widely accepted that they rely on featural processing for recognition (Mondloch, Le Grand, & Maurer, 2002; Pellicano & Rhodes, 2003).

Likewise, face processing “experts” – those who are accustomed to recognizing faces (i.e., older children and adults) – predominantly use holistic processing to recognize a
face (Michel et al., 2006; Richler et al., 2011). Interestingly, some evidence indicates that by 4 to 6 years of age, children show a composite face effect in a similar manner to adults (de Heering, Houthuys, & Rossion, 2007; Mondloch et al., 2007). Recent studies also demonstrate that holistic face processing is predictive of recognition accuracy (Richler et al., 2011; Wang, Li, Fang, Tian, & Liu, 2012), supporting the notion of advanced face processing abilities in adults. There is support that adults employ greater holistic processing when recognizing own-race faces compared to other-race faces, perhaps due to greater experience with own-race faces (Harrison et al., 2011; Hayward, Crookes, & Rhodes, 2013; Michel, Rossion, Han, Chung, & Caldara, 2006). In addition, those who experience deficits in face recognition abilities have impaired holistic processing (Ramon, Busigny, & Rossion, 2010; Teunisse & de Gelder, 2003), suggesting that face processing expertise relies on holistic rather than featural processing. This evidence suggests that children’s immature face processing system is due to their primarily featural nature of face recognition.

Motion provides us with information to perceive and interpret the world around us. It is fundamental that humans can identify what is and is not animate, as this paves the way for appropriate action and human interaction; for example, a ball being thrown toward you may prompt you to avoid being hit, and someone turning their head toward you may prompt a conversation. Because of this, the human face is one of the most important moving objects that we encounter every day. Face motion can be classified into two types: elastic and rigid. Elastic face motion refers to brief transformations in facial musculature (e.g., expressions, eye gaze changes), while rigid face motion refers to changes in face orientation, while its structure remains unchanged (e.g., head-turning,
nodding). Previous studies have found that both types of face motion influenced face recognition (Knight & Johnston, 1997; Pike, Kemp, Towell, & Phillips, 1997).

These movements are visually overt, suggesting that they may provide an observer with distinctive information. The point-light technique (Johansson, 1973) was used in early face motion studies to isolate information provided by facial motion from that of static information. These studies have found that facial motion contains identifying information such as sex, age, facial expression, and identity itself (Bassili, 1978; Berry, 1990, Bruce & Valentine, 1988). Furthermore, studies have shown that observers can accurately identify gender, identity, and kinship solely from facial movement patterns, supporting that face motion itself contains identifying information (Hill & Johnston, 2001; Knappmeyer, Thornton, & Bülthoff, 2003). Recently, researchers have begun examining the role of facial motion on recognition to expand the existing knowledge about face perception. Studies have consistently demonstrated that facial motion enhances recognition abilities for both familiar faces and new faces (e.g., Lander & Bruce, 2000, 2003). Based on this evidence, researchers have proposed two mechanisms regarding facial motion’s beneficial effect on recognition; one suggests that, in addition to static facial information, motion provides the observer with distinctive characteristic information (*supplementary information hypothesis*). This mechanism assumes some experience with a particular individual’s facial motion, and is imperative in learning and recognizing facial motion. The other mechanism suggests that motion provides a flexible and robust three-dimensional face representation for learning new faces (*representation enhancement hypothesis*: O’Toole, Roark, & Abdy, 2002; O’Toole & Roark, 2010). The following studies reveal how facial motion enhances
Knight and Johnston (1997) demonstrated facilitative effects of facial motion on recognition when they presented observers with images and videos of famous people displayed in negative black-white contrast and standard contrast. Recognition performance was higher following the video presentation than the image presentation, but only in the negative contrast condition. This suggests that facial motion information is useful when static information is difficult to access. This finding has been replicated under several non-optimal viewing conditions, such as image blurring, pixilation, and thresholding (Burton, Wilson, Cowan, & Bruce, 1999; Lander, Bruce, & Hill, 2001; Lander, Christie, & Bruce, 1999). Further support for the notion that face motion aids observers when static information is difficult to access comes from studies of individuals with prosopagnosia – a disorder that causes face recognition while other aspects of visual and cognitive processing remain intact (for a review, see Bate, 2013); prosopagnosics were able to use facial motion as a cue to recognize and discriminate among faces based on their movement patterns (Longmore & Tree, 2013; Steede, Tree, & Hole, 2007). These studies suggest that information embedded in facial motion serves a supplementary role to static information for face recognition.

Investigations have also been conducted to determine whether face motion enhances the learning of a novel face. In an early study on the topic, Pike and colleagues (1997) demonstrated that observers recognized faces more accurately when they were learned in rigid motion than from single or multiple frames extracted from the video. Lander and Bruce (2003) used a similar paradigm and found a recognition enhancement following faces presented in elastic motion, perhaps due to the continuity of elastic
motion. Furthermore, observers have been shown to respond to elastic faces faster than static images in a recognition task (Thornton & Kourtzi, 2002). To date, however, no investigation has directly compared rigid and elastic motion on recognition accuracy. Further investigations have demonstrated similar motion facilitation effects using visual search tasks and face matching tasks (Pilz, Bülthoff, & Vuong, 2009; Pilz, Thornton, & Bülthoff, 2006). Interestingly, Pilz et al. (2006) found that face motion improved recognition accuracy even when the test faces were presented at a different angle than the initial learned face; this provides support for a three-dimensional representation enhancement due to motion, and suggests that facial motion’s beneficial effect on face recognition is flexible to changes in face viewpoint.

Further investigation of the beneficial effects of facial motion suggests that naturalness of face motion is necessary for recognition. For example, studies have demonstrated that speed is an important factor – faces presented at normal speed are recognized more accurately than those presented in slow motion (Lander & Bruce, 2004) and sped-up (Lander & Bruce, 2000). In a similar vein, Hill and Johnston (2001) demonstrated that reversing facial motion videos reduced recognition accuracy for gender judgments. Furthermore, faces with natural smiles were shown to enhance recognition compared to faces that displayed a neutral expression that was morphed into a smile (Lander, Chuang, & Wickham, 2006). So, the nature of motion contributes to recognition ability; this suggests that the recognition enhancement effects of facial motion are in part due to the way faces are experienced in real-world settings. Taken together, investigations on the effect of facial motion on recognition suggest that motion provides additional identifying information for familiar faces when static information is difficulty
to access. Evidence also suggests that motion provides rich information for learning new faces, allowing an observer to form an enhanced face representation. Finally, an important factor for recognition is that the perceived motion is natural.

While it is well-established that facial motion leads to improved recognition performance compared to face images, the underlying perceptual mechanisms related to this enhancement remain unclear. Furthermore, although several investigations have examined face processing strategies used for recognition, relatively few studies have examined the influence of face motion information on featural and holistic processing. Does facial motion lead to greater holistic processing or featural processing? Some studies have provided indirect evidence to this question using the previously mentioned face-inversion paradigm. For example, Thornton, Mullins, and Banahan (2011) demonstrated that inversion did not affect gender judgments for static faces, but observers needed significantly more time to judge moving faces when they were presented both upright and inverted. This would suggest that motion facilitated holistic processing; however, other investigations found that inversion did not affect gender and identity judgments (Hill & Johnston, 2001; Knappmeyer et al., 2003). As previously mentioned, inversion was found to impair both holistic and featural processing, and thereafter, the composite face effect task was found to be a more reliable and direct measure of holistic processing.

Recent investigations have used the composite face effect paradigm to determine whether facial motion facilitates holistic or featural processing. Xiao, Quinn, Ge, and Lee (2012) presented novel faces in rigid motion and static conditions. Both conditions were comprised of face images from multiple viewpoints; in the rigid condition, they were
presented sequentially as to emulate natural head rotation movement, while in the static condition, the images were presented with intervals between images or at random, so that motion could not be perceived. At test, participants were told to indicate if the upper half of the composite face resembled the target face, while ignoring the irrelevant lower face half. If the composite face effect were similar for moving and static face conditions, that would indicate that rigid motion does not lead to differential face processing strategies.

Interestingly, the composite face effect was much smaller following rigid motion than static images, implying that rigid motion facilitated featural processing upon recognition. This finding was replicated in a study using elastic motion (Xiao, Quinn, Ge, & Lee, 2013a). Observers learned frontal-view faces that were presenting chewing and blinking movements, or were a still taken from the motion video, and then judged whether the upper half of the composite face belonged to the target face. Not only did elastic motion lead to better recognition, it elicited a smaller composite face effect than the static image condition, similarly augmenting processing of facial features. An additional experiment showed similar effects when participants were asked to recognize the lower face half, which was the target face from the learning phase. This suggests that elastic motion enables observers to adapt to the task requirement and more easily ignore the irrelevant face half.

Further support for the differential processing of face motion is supported by recent studies on eye gaze. Researchers used a design identical to Xiao et al. (2013a) and examined data collected from a high-frequency eye tracker to understand whether higher-order cognitive functioning can account for the enhancement effects of facial movement. The authors found that faces presented in elastic motion led to a significantly longer
duration of visual fixations than static faces (Xiao, Quinn, Wang, Fu, & Lee, 2014). This suggests that elastic movement captures more attention than face images, enabling observers to encode more facial information in moving faces. Furthermore, as observers’ looking time for the upper face half increased, their index of featural processing increased during recognition of aligned composite faces, but not misaligned faces. This indicates that fixating on the upper face half reduced interference of the irrelevant lower face half of aligned composite faces. Importantly, Xiao et al. (2014) also found that elastic motion did not elicit longer fixation time than static faces among individual features; this suggests that motion did not necessarily “pull” observers’ attention toward features. The results from this study indicate that facial motion enhances featural processing by influencing eye movement patterns. Taken together, these studies demonstrate that facial movement leads to a face processing mode that is distinct from that observed in a number of investigations that use static face images. While the evidence here indicates that facial motion promotes featural processing in adults, it is important to address that one recent study on emotion judgment found no difference in the size of the composite face effect between moving and static face presentations. This implies that facial movement can increase the flexibility of face processing under specific task requirements (Chiller-Glaus, Schwaninger, Hofer, Kleiner, & Knappmeyer, 2011).

To better understand the developmental trajectory of facial motion’s influence on processing, Xiao, Quinn, Ge, and Lee (2013d) used the composite face effect task to examine the effects of elastic and rigid motion in children (8-year-olds), adolescents (12-year-olds), and adults (18- to 20-year-olds). The design of the elastic motion experiment was similar to Xiao et al. (2013a), and the design of the rigid motion experiment was
similar to Xiao et al. (2012). The results showed that elastic motion elicited a smaller composite face effect than static images in adolescents and adults, but not children. Rigid motion yielded a smaller composite face effect than static images in children, adolescents, and adults. This suggests that rigid motion leads to featural processing by 8 years of age, while this effect for elastic motion emerges after 8 years. It is important to note that the use of featural processing does not necessarily suggest immature face processing expertise; rather, the ability to use featural processing reflects that motion was able to optimize face recognition.

Prior research suggests that infants are sensitive to facial movements (Ichikawa, Kanazawa, & Yamacuchi, 2011). For example, infants as young as 4 to 8 months were able to differentiate between two models’ idiosyncratic facial movements (Spencer, O’Brien, Johnston, & Hill, 2006). This facial movement sensitivity was supported in a study in which 8-month-old infants viewed schematic faces wherein three black circles representing the eyes and mouth were transformed vertically to represent natural closing and opening, and horizontally to represent biologically-impossible face motion. Infants gazed longer at the unnatural face schematic, signifying a preference for novel stimuli; this suggests that by 8 months of age, infants have a conceptual understanding of natural facial movement (Ichikawa, et al., 2011). Furthermore, Xiao, Quinn, Wheeler, Pascalis, and Lee (2014) found that infants as young as 3 to 5 months of age exhibited a left visual field (LVF) bias in their eye movement patterns for the lower half of naturally-moving face stimuli, while infants ages 6 to 9 months experienced a LVF bias in the whole face area; this suggests that increased experience with faces stimulates a neural response in the right hemisphere, leading to a more robust LVF bias. Investigations have also found that
face motion facilitates recognition just a few hours after birth, and persists throughout the first year of life (Bulf & Turati, 2010; Otsuka et al., 2009). Interestingly, newborns less than 100 hours old were able to recognize a face displayed at a different viewpoint when the face was initially presented in rigid motion, but not static (Bulf & Turati, 2010). Despite these beneficial effects of motion on recognition in infancy, there is some evidence showing that facial movements impair subsequent recognition, suggesting that face motion may distract infants’ processing abilities (Bahrick, Gogate, & Ruiz, 2002; Bahrick & Newell, 2008; Coulon, Guellaï, & Streri, 2011).

To account for these mixed findings regarding facial movement’s beneficial effect in infancy, Xiao et al. (2013b) examined the influence of facial movement on recognition at 3, 6, and 9 months of age. Eye gaze patterns were also recorded. Infants were familiarized with chewing and blinking movements or a static face image, followed by a pair of faces. One of the faces was shown previously in the familiarization phase, and the other one was a novel face. Xiao et al. (2013b) demonstrated that, with increased age, infants fixated longer on the mouth region of the moving face over the static face, and less at the eye region of the moving face than the static face. Furthermore, face motion elicited more fixation shifts among features than static faces, while the center of the face attracted more attention in static faces. The authors also found a significant positive correlation between the number of fixations in the dynamic condition and recognition accuracy at 6 and 9 months, but not 3 months. This suggests that, by 6 months of age, a beneficial role for elastic face motion on recognition accuracy is modulated by eye movements. Additionally, this study suggests that facial movement captures attention toward moving face parts beginning in infancy. Taken together, these studies demonstrate
that infants’ sensitivity to facial movements may emerge by 3 months of age and
develops into the ability to recognize faces based on face motion characteristics.

To better understand the developmental trajectory of facial motion’s influence on
processing, Xiao, Quinn, Ge, and Lee (2013d) used the composite face effect task to
examine the effects of elastic and rigid motion in children (8-year-olds), adolescents (12-
year-olds), and adults (18- to 20-year-olds). The design of the elastic motion experiment
was similar to Xiao et al. (2013a), and the design of the rigid motion experiment was
similar to Xiao et al. (2012). The results showed that elastic motion elicited a smaller
composite face effect than static images in adolescents and adults, but not children. Rigid
motion yielded a smaller composite face effect than static images in children,
adolescents, and adults. This suggests that rigid motion leads to featural processing by 8
years of age, while this effect for elastic motion emerges after 8 years. It is important to
note that the use of featural processing does not necessarily suggest immature face
processing expertise; rather, the ability to use featural processing reflects that motion was
able to optimize face recognition.

Does face motion influence processing strategy and eye movement patterns in
children? In the present study, we examined the influence of elastic face motion on
recognition and face processing strategies in 4- to 6-year-old children using the
composite face effect paradigm to gain an understanding of facial motion’s effects on
processing in early childhood. Furthermore, we examined eye movement patterns to
study possible higher-order cognitive processes that may influence dynamic face
processing. Participants were familiarized with a frontal view of a young adult dynamic
or static target face. The dynamic face was blinking and chewing silently, while the static
face was an image extracted from the dynamic face video. Following the familiarization phase, the test phase depicted a pair of aligned or misaligned composite faces. Of the pair, one upper face half belonged to the familiarized face, while the other upper face half belonged to a novel face. Both lower face halves belonged to a separate novel face. The participants’ task was to choose which upper face half belonged to the face previously displayed. The composite face effect index was computed by using response accuracy differences between the misaligned and aligned conditions, similar to previous studies (Mondloch & Maurer, 2008; Xiao et al., 2013a; Young et al., 1987). If participants performed better on misaligned trials than aligned trials, a composite face effect should be observed, indicating that the target face was processed holistically; if there is no difference in performance between misaligned and aligned trials, a composite face effect will not be observed, indicating that the target face was processed featurally. To examine how elastic face motion influences face processing strategy, the composite face effect for dynamic target trials will be compared to that of the static target trials. A larger composite face effect in the dynamic condition would indicate that elastic face motion facilitates holistic processing, while a smaller composite face effect in the dynamic condition indicates that elastic motion facilitates featural processing. If the composite face effect indices are similar, that indicates that elastic face motion does not facilitate processing strategy any differently than observing static face images.

We used eye-tracking software to record observers’ eye movement patterns in the familiarization phase to examine whether face motion elicited different fixation durations among facial features. This will shed light on the development of face motion gaze at a critical period in which holistic face processing abilities emerge. Furthermore, we
explored whether individual fixation durations were related to recognition accuracy. We created the following AOIs (areas of interest) on each video and image used in the familiarization phase of the experiment: left eye, right eye, nose, mouth, upper half, lower half, and whole face. We then converted fixation data to proportional looking time within each AOI relative to the whole face looking time, as examined in previous studies (e.g., Wheeler et al., 2011); this calculated fixation duration for each AOI as a percentage so they may be compared.

**Method**

*Participants*

The final sample was comprised of 79 Chinese children ages 4 to 6 years old ($M = 62.7\text{ months, } SD = 6.8$). An additional 16 participants were removed from the analysis – 10 of which exhibited response bias (either consistently chose one response option or alternated between both options), 3 had low overall accuracy (less than 50% accurate), and 3 had both response bias and low accuracy. All children had normal or corrected-to-normal vision. Participants were not familiar with any of the face stimuli in the experiment. Informed consent was obtained from the child’s parent prior to their participation in the study.

*Stimuli and Materials*

The individuals who were recorded are referred to as the “models”, and the individuals tested in the experiment are referred to as the “participants”. There were 20 models (10 male, 10 female), all of whom were Chinese and 19-24 years of age. To create the stimuli for the experiment, the models were instructed to pose facing the camera with a neutral facial expression, and avoid making head movements, and silently
produce chewing and blinking motions. Four 2000 ms videos were recorded for each model, used in the dynamic condition of the familiarization phase; one still image was extracted from each video, used in the static condition of the familiarization phase (see Figure 1). In this phase, the dynamic or static face was displayed in the center of the image for 2000 ms. In the test phase, two composite faces were displayed side-by-side.

The target composite face was created by merging the top half of the target face previously displayed in the familiarization phase with the bottom half of a different face not used in the experiment. In the aligned condition, the face halves were adjusted so that they appeared to create one whole, complete face. In the misaligned condition, the bottom face half was moved to the right by approximately half the width of the face. The foil composite face was created in the same manner as the target composite face except that both halves were not previously seen (see Figure 2). For each pair of composite faces, the external features of the top half (i.e., hair and ears) were identical so that participants’ recognition was based on internal features. Furthermore, the bottom face halves for each pair of composite faces were identical.

Throughout the experiment, an eye tracker system was used to record participants’ eye movement patterns in the familiarization phase. The EyeLink 2000 with a sampling rate of 500 Hz was positioned underneath the display, approximately 60 cm from the participant. A calibration procedure included a small cartoon figure displayed at 5 positions: center, the middle of top, bottom, left, and right edges of the screen. The calibration procedure was terminated once the child was calibrated on all five positions.

Procedure
Children were tested individually in a quiet and brightly lit room. They were seated approximately 60 cm from the display, with a visual angle of approximately 44°. ExperimentBuilder software was used to display the experiment on a 19-inch monitor with a resolution of 1024×768 pixels and a frame rate of 60 Hz. All videos and images were presented on a black background with a size of 640×768 pixels.

Participants completed a child-friendly version of the CFE paradigm. Each experimental trial consisted of two phases – the familiarization phase and the test phase. Once the trial began, a fixation cross was displayed in the center of the screen for 500 ms. The familiarization phase began after the offset of the fixation. In the dynamic condition, a moving face was displayed for 2000 ms. Participants were instructed to remember the top half of the face shown in the familiarization phase. After this phase, a visual mask was presented for 500 ms to remove possible afterimage effects. The test phase began following the visual mask. In this phase, the two aligned or misaligned composite faces were shown side by side; the face on the left was surrounded by a green box, while the face on the right was surrounded by a red box. Faces were displayed until the participant made a response. Participants indicated which face resembled the one previously shown by saying out loud “green” or red”, and the experimenter entered this response. In the static aligned and misaligned conditions, a face image was extracted from the face motion video, and was displayed for 2000 ms. All participants completed 80 experimental trials, separated by a short break after the first 40 trials. The experiment was equally divided into 2 (target face type: dynamic and static) × 2 (composite face alignment: aligned and misaligned) = 4 conditions. All trials were presented in a random order, and each condition was distributed proportionately across the 80 trials.
**Data Analysis**

The measurement used from the CFE task was participants’ response accuracy. As previously mentioned, studies in the face processing literature examined only accuracy data from the composite face trials in the test phase; this method of examining face processing type is considered to be reliable (Richler, et al., 2011). Two composite face effect indices were calculated by subtracting aligned condition recognition accuracy from misaligned condition recognition accuracy for both dynamic and static trials.

Matlab software was used to create AOI (area of interest) coordinates on each image and video for all faces presented in the familiarization phase of the experiment. Limited head movement allowed for individual facial features to be contained inside the AOIs throughout each video. The following AOIs were created: left eye, right eye, nose, mouth, upper face half, lower face half, and whole face (see Figure 3). It should be noted that the mouth AOI was created considerably larger than in previous investigations using static faces to account for the chewing motion. As previously mentioned, fixation data were converted to proportional looking time within each AOI relative to the total whole face looking time, producing values as fixation duration percentages. Trials in which the whole face looking time was less than 100 ms were removed from the analysis. Furthermore, it is important to note that several participants reported experiencing fatigue after the second half of the experiment; therefore, only the first 40 of 80 total trials were included in the final analysis. Within these 40 trials, each condition was distributed proportionately. Finally, preliminary analyses revealed that participants’ age in months was not related to any variables of interest, suggesting that the following findings were similar among participants regardless of age.
Results

To investigate the effect of facial motion and composite face alignment on recognition, accuracy scores were submitted to a two-way repeated measures analysis of variance (RM-ANOVA) with the within-subjects factors target face type (dynamic, static) and composite face alignment (misaligned, aligned). There was no difference in response accuracy between dynamic and static presentations ($M_{Dynamic} = .73$, $M_{Static} = .72$), $F(1, 78) = 1.88$, $p = .174$, $\eta^2_p = .024$. This suggests that facial motion does not enhance recognition among young children. Additionally, no difference in response accuracy was found between misaligned and aligned composite face conditions (see Figure 4; $M_{Misaligned} = .72$, $M_{Aligned} = .74$), $F(1, 78) = 2.16$, $p = .146$. Because misaligned faces were not better recognized than aligned faces, we can infer that the irrelevant lower face half did not interfere with recognition of the upper half when the halves formed a whole face. This indicates an absence of the composite face effect, suggesting that these faces were not processed holistically, and therefore assumes that featural processing was the primary method used for recognition regardless of target face type.

We then examined the composite face effect to determine whether there was a difference in face processing strategy between dynamic and static presentations. The composite face effect was calculated for the dynamic and static conditions by subtracting response accuracy of the aligned condition from the misaligned condition (i.e., $dynamic\ CFE = dynamic\ misaligned - dynamic\ aligned$; $static\ CFE = static\ misaligned - static\ aligned$). A paired-samples t-test revealed no difference in CFE scores ($M_{Dynamic} = .003$, $M_{Static} = -.043$), $t(78) = 1.58$, $p = .119$, suggesting that observing facial motion does not facilitate face processing type (i.e., holistic, featural) any differently than a static image.
Next, we investigated whether dynamic and static face presentations yield different fixation patterns among the eyes, nose, and mouth. Proportional looking time was submitted to a two-way RM-ANOVA with the within-subjects factors target face type (dynamic, static) and facial feature (eyes, nose, mouth). Results revealed a main effect of target face type, $F(1, 78) = 14.51, p < .001, \quad \frac{\nu}{\sigma} = .157$. Fixation duration among the eyes, nose, and mouth were significantly longer in the dynamic condition ($M_{\text{Dynamic}} = .25$) compared to the static condition ($M_{\text{Static}} = .24$), suggesting that these features captured more attention when the target face was moving. As shown in Figure 5, there was a significant interaction between target face type and facial feature, $F(2, 156) = 43.289, p < .001, \quad \frac{\nu}{\sigma} = .357$. Follow-up $t$-tests found that fixation duration for eyes was significantly shorter in the dynamic condition ($M_{\text{Dynamic}} = .27$) than the static condition ($M_{\text{Static}} = .32$), $t(78) = -4.71, p < .001$. By contrast, fixation duration for mouth was significantly longer in the dynamic condition ($M_{\text{Dynamic}} = .21$) than the static condition ($M_{\text{Static}} = .11$), $t(78) = 9.87, p < .001$. Follow-up post-hoc comparisons (LSD, $\alpha = .05$) indicated that the eyes captured more fixation time than the mouth for dynamic faces ($M_{\text{Eyes}} = .27, M_{\text{Mouth}} = .21$), $t(78) = 2.23, p = .029$, and for static faces ($M_{\text{Eyes}} = .32, M_{\text{Mouth}} = .11$), $t(78) = 9.23, p < .001$. These results suggest that, while the eyes were fixated on longer than the mouth for both dynamic and static faces, face motion elicited significantly higher mouth fixation time and less eye fixation time than static images, indicating that motion attracted observers’ attention toward the mouth area.

We conducted a follow-up test to determine if face recognition accuracy and composite face effect indices can be predicted by observers’ fixation patterns for dynamic and static faces. Correlation analyses revealed that fixation duration on the mouth was
influential for recognition. For dynamic faces, mouth fixation was moderately negatively related to recognition accuracy, $r(79) = -0.217$, and for static faces, mouth fixation was significantly negatively related to recognition accuracy, $r(79) = -0.254$, $p = .024$. These results suggest that observers performed worse on the task the longer they fixated on the mouth area. No relationships were found among fixation patterns and composite effect scores, suggesting that no specific area of the face influenced the primary face processing strategy used.

Discussion

In the present study, we examined the effect of elastic face motion on face processing strategies and observers’ eye movement patterns in early childhood. Specifically, we investigated: 1) whether elastic face motion enhances recognition of novel faces; 2) whether elastic face motion facilitates holistic or featural processing; and 3) whether elastic face motion elicits a difference in fixation duration among facial features. From this investigation, we obtained four major findings. First, face motion did not enhance face recognition in early childhood. Second, composite face effect scores for both moving and static faces were not significant, suggesting that both target face types were processed featurally. Third, there was no significant difference in CFE scores between moving and static faces, suggesting that observing dynamic and static faces led to featural processing at a similar level. Third, eye-tracking data revealed that the mouth area received longer fixation times than the eyes when the face was moving, while static faces elicited higher fixation duration for the eyes than the mouth. This suggests that moving face parts (i.e., a chewing mouth) capture more attention than still faces in young children.
To our knowledge, the current study provides the first direct evidence about the influence of moving faces on face processing strategy and eye movement patterns in young children. Our finding that elastic face motion did not enhance recognition abilities does not support the representation enhancement hypothesis for learning new faces (O’Toole et al., 2002), suggesting that motion does not enhance the three-dimensional face representation in young children. It should be noted that the supplementary information hypothesis does not apply in this investigation since the faces were not previously known to the observer, and therefore we cannot suggest that face motion did not provide additional identifying information. Importantly, this finding reflects developmental differences in the role of motion on face learning; it contrasts with the face motion enhancement effect found in adults (Lander & Bruce, 2003, 2004; Pike et al., 1997; Pilz et al., 2006; Xiao et al., 2012, 2013a). This suggests that young children do not yet possess the ability to use face motion to optimize face processing for recognition, which is developed sometime after six years of age.

It is important to note that the absence of an enhancement effect may be due to the nature of the test faces; only the upper face half of the target composite face resembled the one previously seen. The upper half of the moving face was blinking rather infrequently (approximately 1-2 blinks per 2-second video clip), while the lower half of the moving face contained exaggerated chewing motions throughout the 2-second video clip. Clearly, the lower half of the face contained substantially more motion information compared to the upper half. This would suggest that the requirement of the CFE task limited observers’ use of face motion for recognition.

We compared the composite face effect for dynamic faces to static faces and
found no significant difference in scores, indicating that moving faces led to featural processing comparable to static faces. Similar findings have been revealed in studies with adults (Xiao et al., 2012; 2013) and studies that have indirectly assessed face processing strategy using inversion (Hill & Johnston, 2001; Knappmeyer et al., 2003). This result is not surprising, as evidence suggests that young children rely on processing facial features for recognition of face images (McKone, 2008; Mondloch et al., 2003; Pellicano & Rhodes, 2003), though other research suggests that young children show a similar composite face effect to adults (de Heering et al., 2007; Mondloch et al., 2007). These mixed findings may be due to a critical period in the development of holistic face processing during early childhood. In terms of the face processing strategy used when faces are learned in motion, however, Xiao et al. (2013d) found that holistic processing is prevalent by 8 years of age. This suggests that holistic processing of moving faces develops sometime after 4 to 6 years, and before 8 years. From this, we can infer that face motion processing expertise develops slower than that of static faces.

There exists at least two nonmutually-exclusive possibilities for the comparatively gradual development of face motion processing. First, elastic face motion provides more information to an observer than a face image. It is possible that young children are unable to efficiently process this information effectively for recognition. The second possibility is that elastic face motion draws attention to individual facial features (Lander, Hill, Kamachi, & Vatikiotis-Bateson, 2007). This is supported by our finding that dynamic face presentations elicited longer fixation times for facial features than static face presentations. Specifically, motion drew attention to the mouth, which was exhibiting stronger motion cues relative to the rest of the face. Previous studies have shown that
featural processing becomes dominant in adults when faces possess distinctive facial features (e.g., Moscovitch, Winocur, & Behrmann, 1997). In a similar vein, elastic motion may cause the moving facial features to appear more prominent than those that remain static (Xiao et al., 2013a). Eye-tracking patterns revealed important information about the developmental course of eye movement patterns for elastic face motion and face images, speaking to the possibility of facial motion capturing attention at specific features. In adults, elastic motion did not elicit longer fixation durations than static images among individual facial features (Xiao et al., 2014). So, unlike adults, elastic motion pulled young children observers’ attention toward moving facial features. Interestingly, the young children in this study exhibited eye movement patterns similar to infants, who fixated longer on the mouth of moving compared to static faces, and less at the eyes of moving compared to static faces (Xiao et al., 2013b). This suggests that young children are comparably susceptible to the attention-capturing effects of face motion as infants. Because adults show no differences in fixation time among moving face parts, this suggests the role of an expert face motion processing system whereby adults take into account the whole face and are not distracted by moving facial features. Perhaps this may account for the beneficial effect of face motion in adults.

Taken together with recent investigations about face motion processing, the present findings reveal a developmental trajectory of facial motion’s influence on processing strategy and eye movement patterns. The present study also revealed important implications for our understanding about the mechanisms behind face processing in the real world. As previously mentioned, the faces we encounter in real-world settings are constantly moving, while most of our current knowledge about face
processing comes from investigations using static face stimuli. Whereas individuals primarily use holistic processing when observing face images, they rely on featural processing when observing moving faces (Xiao et al., 2012, 2013a, Xiao et al., 2013d). This heavily implies that natural face processing strategy may not be primarily holistic, in sheer contrast to what we know about from investigations using static face stimuli. Therefore, it is important for researchers to evaluate and reinterpret evidence about face processing to include the processing of face motion.

Given the prominence of real-world face processing, it should be noted that the present study tested recognition using static composite faces. Ideally, researchers should use a design that presents both the familiarization and test phases in static and in motion; however, it is difficult to produce moving composite faces since these faces require matching of two face halves. For static faces, image properties such as facial feature shape and skin colour can be modified easily, while for dynamic faces, it may be difficult to match face movement patterns. Unless the movements of the two separate face halves were identical and moved in sync, the aligned condition would not be perceived as a whole face. Regardless of this obstacle, by using static composite test faces, comparisons can be drawn among previous studies that have used similar stimuli. As previously mentioned, most of the facial movement information was contained around the mouth region, attending observers toward the lower face half. Perhaps face motion would have enhanced recognition in this task if more motion information was available in the upper face half, or if the task required observers to remember the lower face half. Since fixation duration for the upper half of the face predicted recognition accuracy, this suggests that the requirement of the CFE task limited observers’ use of face motion information.
Finally, due to the nature of the stimuli in the test phase, we cannot make solid conclusions about the effect of elastic face motion on recognition and processing strategies, at least in every context.

In sum, the findings outlined in this investigation demonstrate that children process face motion differently than static face images. Unlike adults, young children have not developed the ability to use face motion information to enhance recognition. Face motion activated featural processing, though no differently than static face images, and moving face parts attracted more attention than when the face was still. Future research can explore different approaches to examining face motion, including the use of methods such as the part-whole paradigm (see Tanaka & Farah, 1993). Perhaps future work will shed light on the developmental course of face motion processing.
References


of specialized processing of own-race faces in infancy. *Infancy, 14*, 263-284. doi:10.1080/15250000902839369


Pascalis, O., de Martin de Vivies, X., Anzures, G., Quinn, P. C., Slater, A. M., Tanaka, J.


Figure 1.
Figure 2.
Figure 3.
Figure 4.
Figure 5.